

Atlas Neural Networks as Self-Characterizing Systems

One intriguing property of Atlas neural networks is that they naturally expose meaningful internal statistics. Rather than treating the network as a largely opaque function approximator, the routing structure and local experts provide a rich source of diagnostic information that can guide both model analysis and model improvement.

Some examples include:

- **Context bit statistics**

- Bit entropy and utilization.
- Pairwise correlations between bits.
- Mutual information with the target.
- Distribution of context activations.

- **Region statistics**

- Visitation frequency.
- Local prediction error.
- Margin distributions.
- Gradient norms.
- Fisher information or other importance measures.

- **Local model statistics**

- Weight norms.
- Numerical conditioning.
- Effective rank.
- Sensitivity to reduced numerical precision.

These statistics have direct engineering value.

- Low-entropy or highly correlated context bits may indicate opportunities to simplify the routing structure.
- Rarely visited regions may be candidates for pruning or merging.
- Regions with consistently high error suggest where additional capacity would be most beneficial.
- Quantization-sensitive regions can retain higher precision while simpler regions are compressed more aggressively.

This creates a closed-loop engineering workflow:

1. Train the network.
2. Collect routing and regional statistics.
3. Identify structural inefficiencies.
4. Adjust the atlas topology or model hyperparameters.
5. Retrain and repeat.

In this sense, Atlas neural networks are **self-characterizing**. Their architecture not only performs inference but also provides interpretable measurements describing how effectively its own partitioning of the input space matches the data. Rather than relying solely on global metrics such as validation accuracy, the model exposes localized diagnostics that can directly inform architecture refinement, compression, quantization, and future hyperparameter choices.